Wolfpack

(Windows NT® Clusters)

Sivaprasad Padisetty

Microsoft Corporation

Today's Agenda

- What is Clustering?
- Project goals
- Architecture overview
  - Obligatory Demo
- Clustering API Overview

What is a Cluster?

- Group of independent systems that appear to be a single system
- Managed as a single system
- Common name space
- Services are “cluster wide”
- Can tolerate component failures
- Components can be added transparently to users

Development Goals

- Extend Windows NT to seamlessly include cluster features
- Easy to use, mass market product
- Availability
- Manageability
- Scalability
- Support for file/print, web Srv, well behaved apps as base features
- API for “cluster aware” products

Windows NT Clusters

Target applications

- Database servers
- E-mail, Groupware, and Productivity Applications
- Transaction Processing Servers
- Enterprise Applications (i.e.; SAP, CA/Unicenter)
- Internet Web Servers
- File and Print Servers

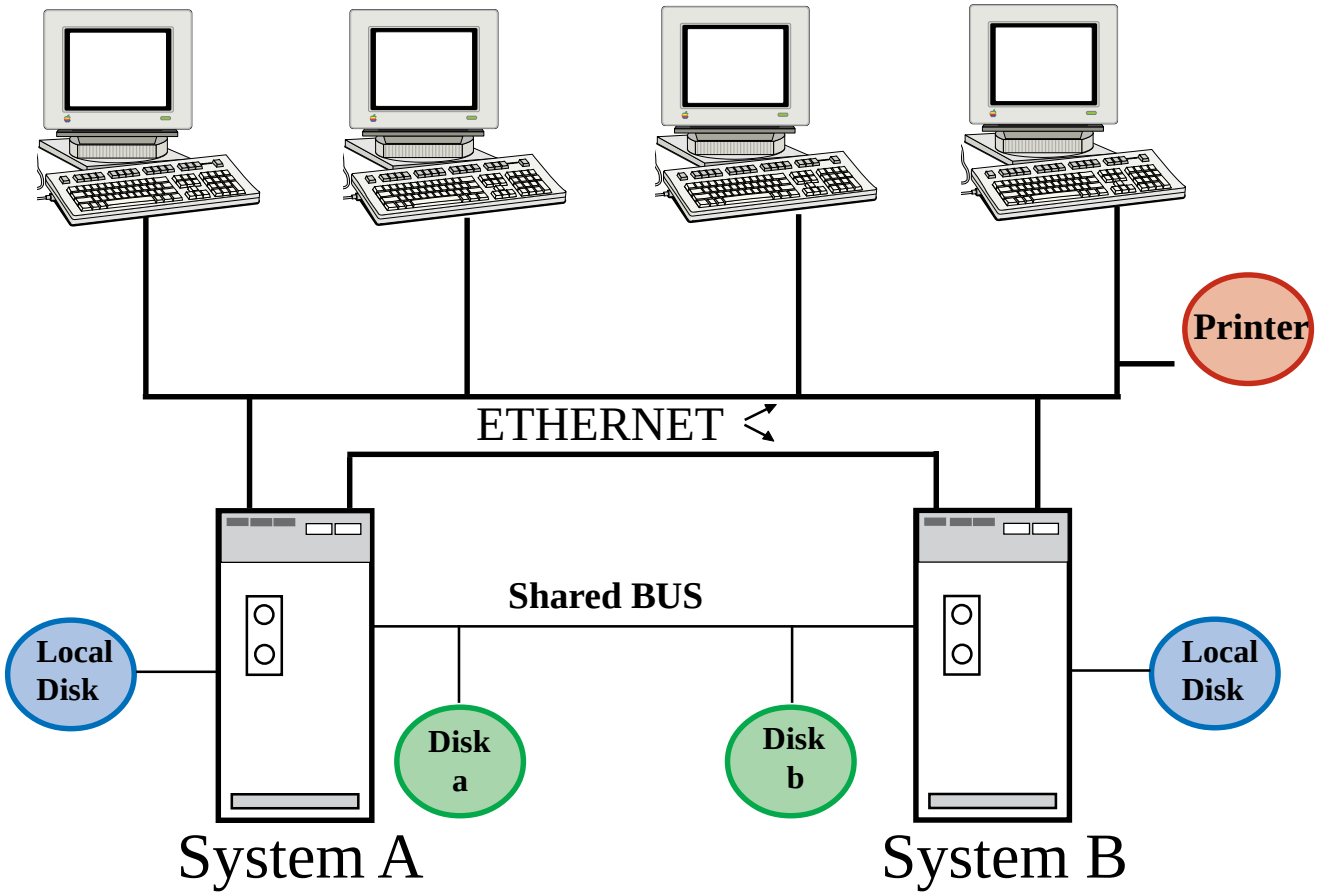
Windows NT Clusters

Phase 1 Features

- Two nodes (design is for N)
  - Both are active, no passive server
- Fail over protection for server applications
- Applications are explicitly registered for support
- Dual-access storage model
  - Both nodes can access storage, one at a time (one node “owns” it)

Windows NT Clusters

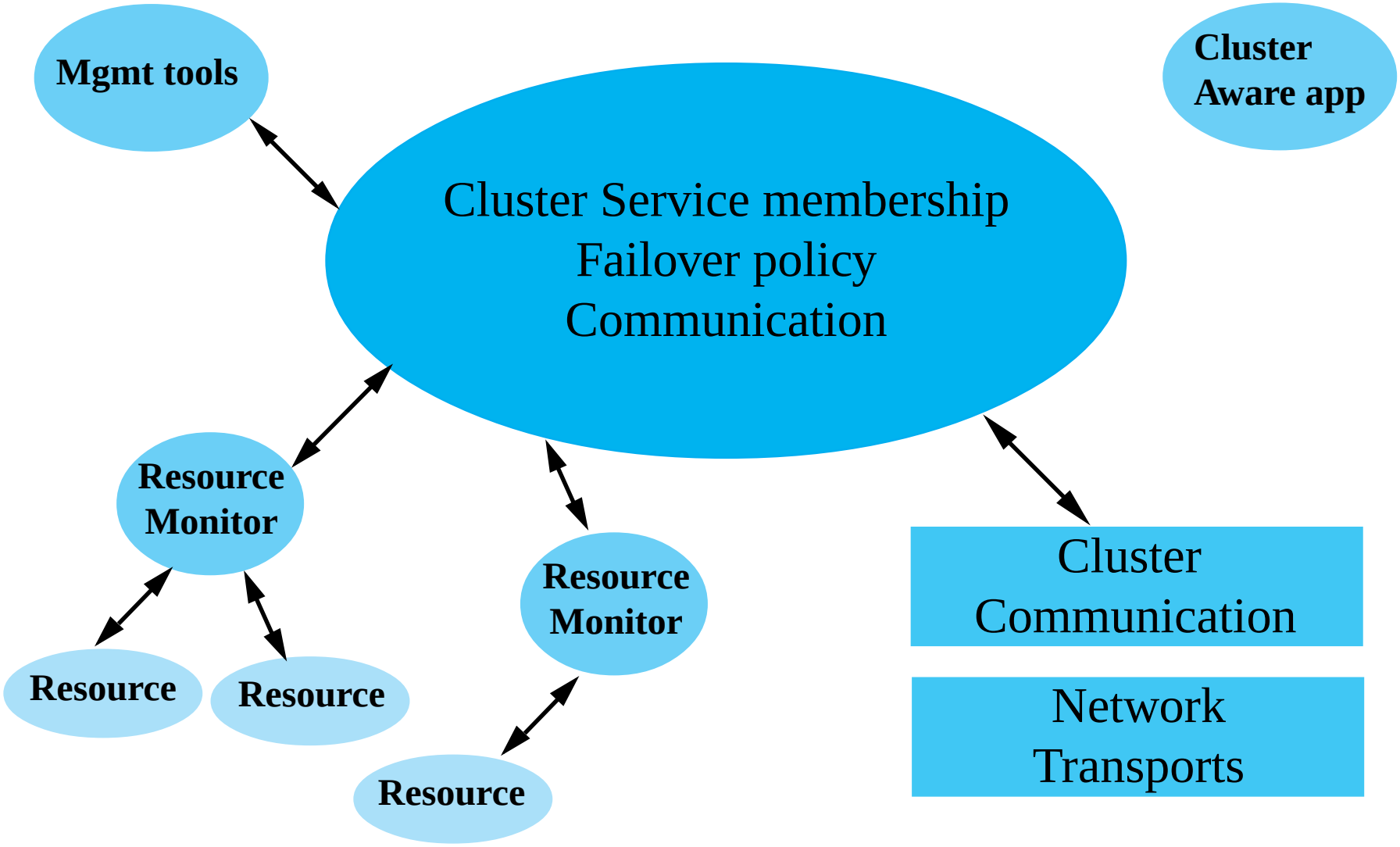
Phase 1 Configuration



Design Decisions

- Separate service rather than modify NT Kernel
  - Does not burden base system
  - Easier to design and debug
  - Slight performance and space cost
- Shared nothing
  - Simplified hardware configuration
  - No shared file system
- Very powerful administrative tools
- Applications must be transactional

Wolfpack Components



Basic Wolfpack Terms

- Cluster – one or more closely-coupled nodes, managed as a single entity
- Node – NT Server running cluster software
- Group – collection of Resources
- Resource – basic entity managed by the cluster

Resources

- Components that provide a service to clients in a client server environment such as:
  - physical disks,
  - processes, databases,
  - IP addresses, etc. …
- Managed by cluster as Opaque objects
- Online on one node at a time
- Can move from one system in the cluster to another system in the cluster

Resource Dependencies

- Resources may depend on other resources
- A resource is brought online after any resources it depends on
- A resource is taken offline before any resources it depends on
- All dependent resources will failover together and exist in the same Group

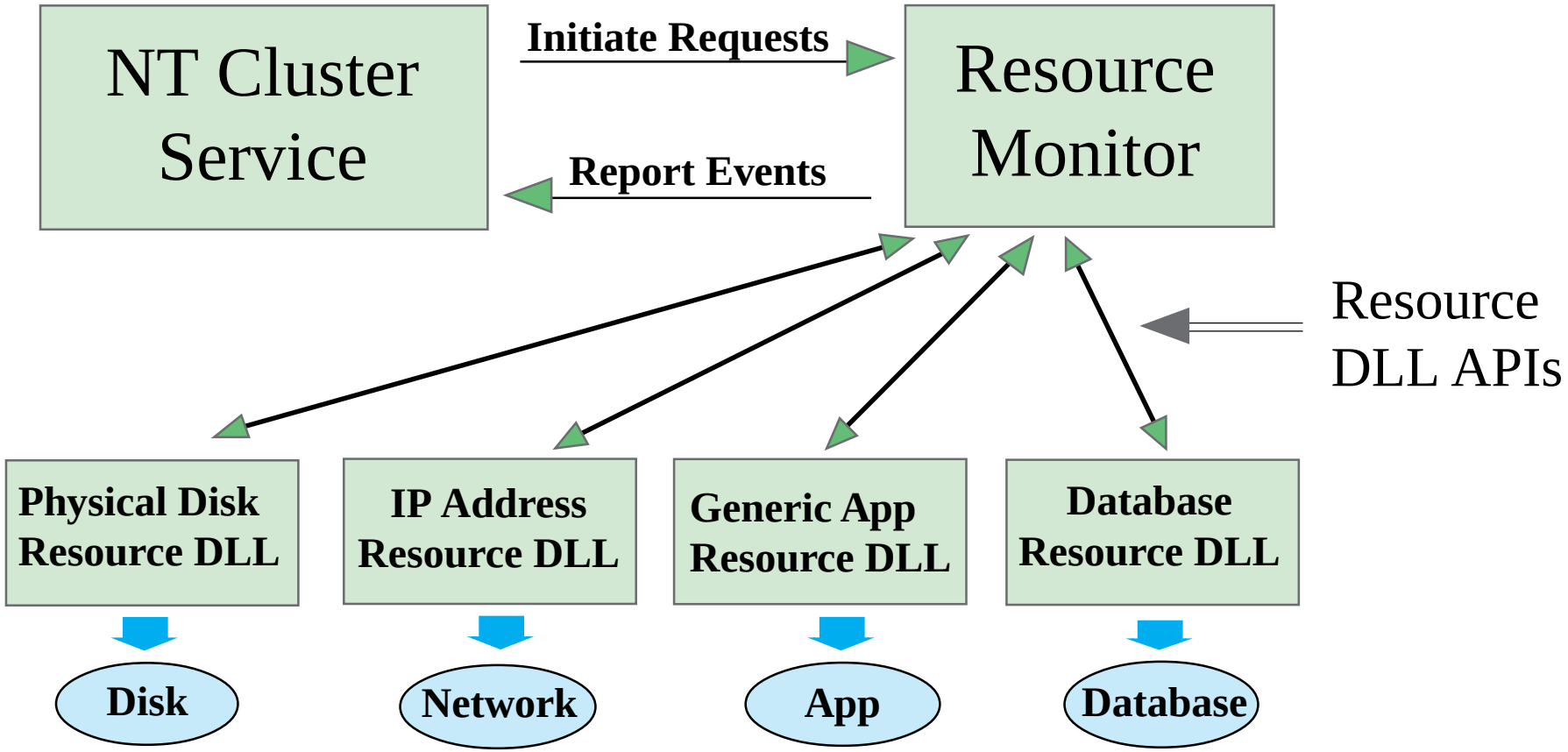
Resource Properties

- Resource Type
- Poll intervals
  - LooksAlive
  - IsAlive
- Private resource data
  - Unique identifier
  - Hardware binding
- Group Membership
- Possible Nodes
- Restart Policy
- Dependencies

Resource DLLs

- Run in a resource monitor process
- Resource specific code that allows cluster software to manage a resource
- Resource DLLs will be provided by Microsoft for base features
- Easily written to create cluster aware applications or devices

Resource DLL APIs



Resource DLL APIs

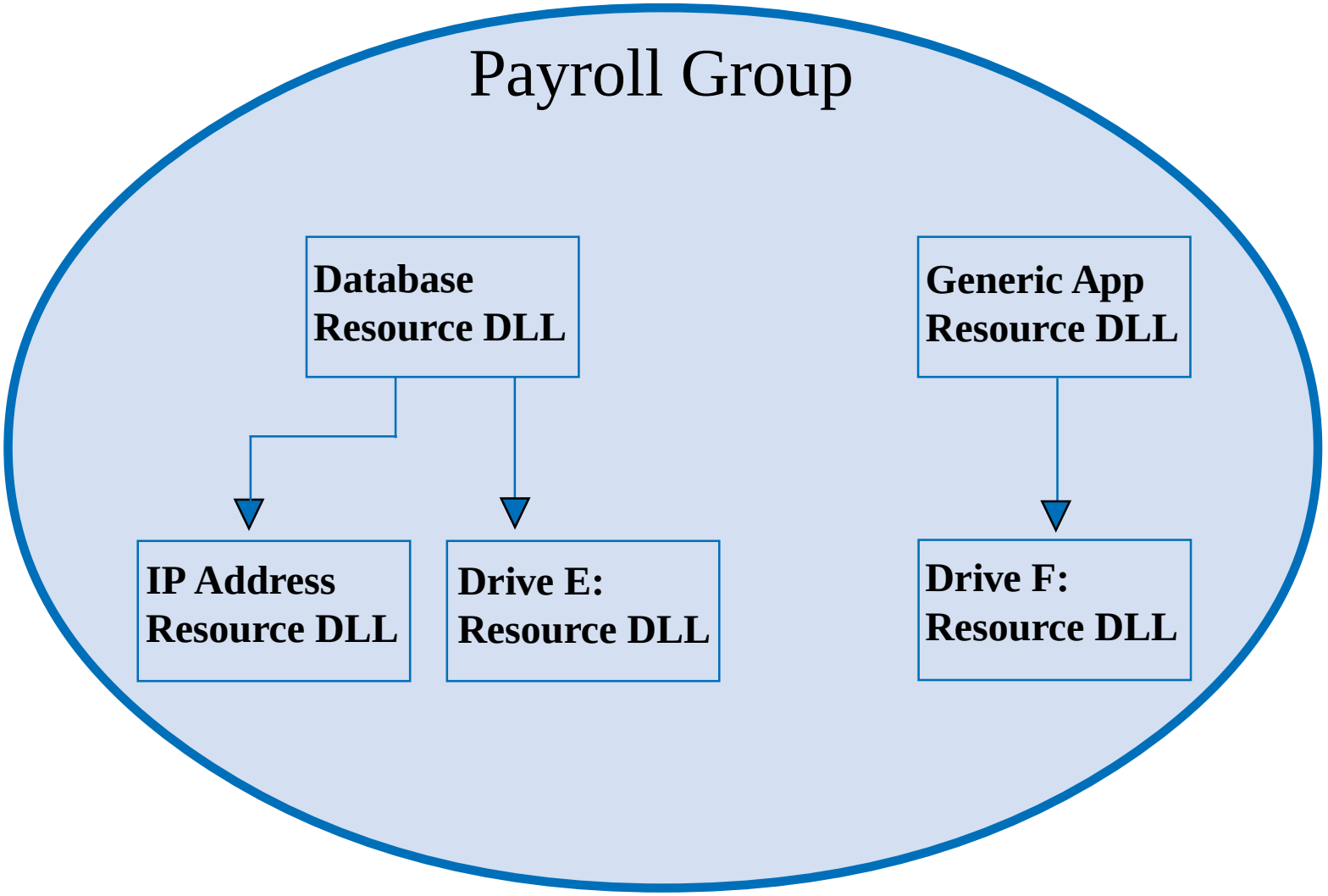
Functions

- Startup

- Open
- Close
- Online
- Offline
- Terminate
- LooksAlive
- IsAlive

- Arbitrate
- Release
- ResourceControl
- ResourceTypeControl
- Create
- Delete

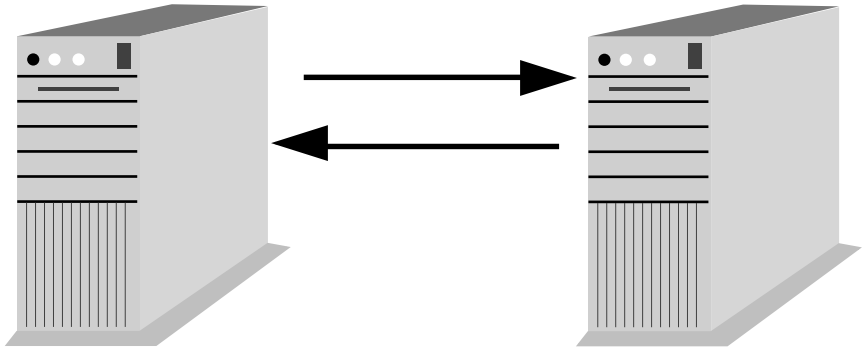
Group Example



Connecting the Pieces

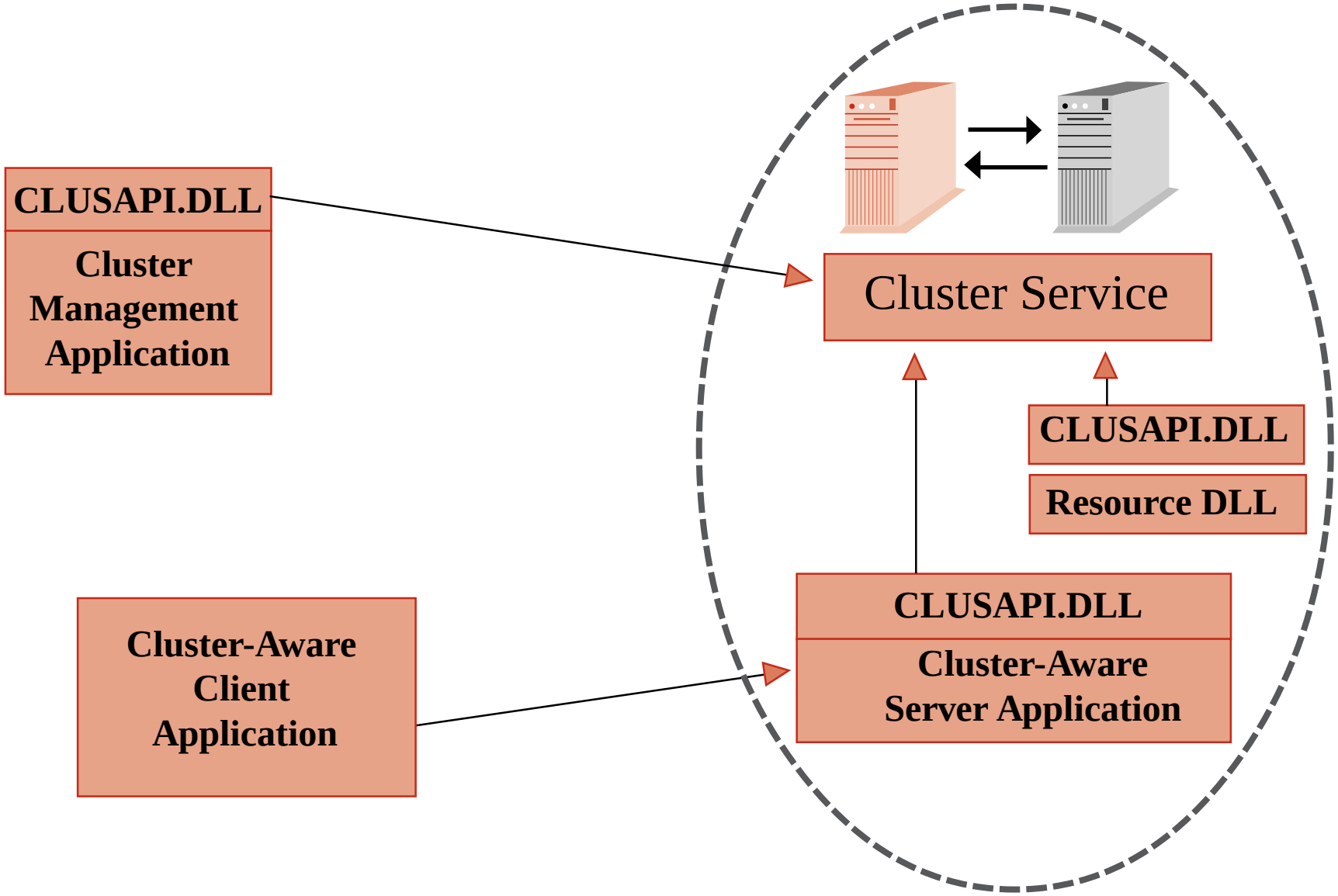
- Node-to-Node communications
- Cluster Service to Resource communications
- Application to Cluster Service communication

Node to Node



- Cluster Services talking to each other
- Keepalive notifications
- Resource state change atomic broadcasts
- Intra-cluster commands

Application to Cluster Service



Cluster Management Tools

- Cluster administrator
  - Monitor and manage the cluster
- Minor modifications to existing tools
  - Performance monitor
    - Add ability to watch entire cluster
  - Disk administrator
    - Add understanding of shared disks
  - Event logger
    - Broadcast events to all nodes

Managing Applications

- Many applications will “just work” without any modi fications
  - Support for generic applications and services
  - Tools configure them into Wolfpack
  - Clients frequently “gracefully” handle server brief outages
- Easy to support other applications
  - Wizard to create a “resource DLL”
  - Wizard to create administrator extensions

CLUSAPI

- Cluster Management
- Cluster Node Management
- Cluster Resource Management
- Cluster Configuration Database

Cluster Management

- Find and communicate with Cluster
- Query/Set Cluster properties
- Enumerate Cluster objects
  - Nodes
  - Groups
  - Resources and Resource Types
- Cluster event notifications
  - Node state and property changes
  - Group state and property changes

Summary

- Wolfpack – cluster features for NT
- First phase concentrates on availability
  - File, print, IIS built-in to phase 1
  - Simple to install and use
  - Improves existing application availability
  - Foundation for future features
- Second phase real distributed apps